

一种基于模糊神经网络加权的多维稀疏模糊推理方法*

刘文远¹ 张庆大¹ 王宝文¹ 石岩²

(燕山大学信息科学与工程学院 秦皇岛 066004)¹ (九州东海大学工程学院信息系统系 日本)²

摘要 在稀疏规则库条件下,经典的插值理论针对一维稀疏规则库提出了各种不同的插值方法,取得了很好很多的经验;但对多维稀疏规则条件的近似推理研究很少,不仅存在着难以保证推理结果的凸性和正规性等问题,而且没有考虑到多维变量之间的联系即对结论的影响权值,造成推理结果的误差性更大。多变量规则的模糊插值推理是插值推理研究的重要方面,为了在多变量稀疏规则条件下得到好的插值推理效果,本文提出了一种基于模糊神经网络加权的多维模糊推理方法,为智能系统中的模糊推理提供了一个十分有用的工具。

关键词 多维稀疏模糊推理,模糊神经网络,权值,相似性

Multidimensional Fuzzy Reasoning Method Based on Weight of Fuzzy Neural Network

LIU Wen-Yuan¹ ZHANG Qing-Da¹ WANG Bao-Wen¹ SHI Yan²

(School of Information Science and Engineering, Yanshan University Qinhuangdao 066004)¹

(School of Information Science, Kyushu Tokai University, Japan)²

Abstract Interpolative reasoning is type of important reasoning approaches under sparse rules. Interpolative reasoning in one dimension has been researched widely, but the research in multi-dimension is lacking and a few existing approaches have some faults. These methods not only cannot guarantee the convexity of result, but also cannot consider the relation between many variables, weight of influenced conclusion. It leads to more error of inferential result. Interpolative reasoning in multi-dimension is an important research aspect of interpolative reasoning, in order to get better conclusion under multidimensional sparse rules condition, we propose a fuzzy multidimensional reasoning method based on weight of fuzzy neural network, which moreover can keep the convexity of the reasoning consequence.

Keywords Multidimensional sparse fuzzy reasoning, Fuzzy neural network, Weight, Similarity

1 引言

在模糊控制、模糊专家系统、模糊决策等智能系统的开发和研究中,模糊推理是其中一个十分重要的组成部分。目前在模糊推理中使用得最多的模糊推理方法是 Zadeh^[9] 提出的合成推理规则(CRI),但在稀疏规则条件下,规则不能够覆盖整个论域,规则之间有“空隙”,如果输入事实落在“空隙”中,则CRI方法不能够适用。为了解决此问题,学者们证明了模糊推理其实就是插值器^[8]。KH插值推理^[1,2]是稀疏规则条件下的重要推理方式,并且运用也越来越广泛。

我们将以 Mizumoto 和 Zimmerman^[11] 提出的西红柿问题为例介绍稀疏模糊规则库。

通常有以下的正常推理:

前提1 如果西红柿是红色的,那么西红柿就熟了。

前提2 如果这个西红柿是红色的,

结论 这个西红柿就熟了。

但是在下面的推理中就有困难了:

前提1 如果西红柿是红色的,那么西红柿就熟了。

前提2 如果西红柿是绿色的,那么西红柿就不熟。

前提3 这个西红柿是黄色的

结论 ???

我们从直觉上可得出这个西红柿是半熟,但是在规则库

中找不到匹配的规则,因此在模糊推理中得不到结论,如图1所示。

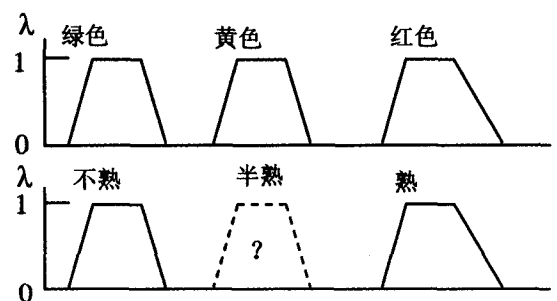


图1 西红柿问题

在稀疏规则条件下,对于单变量模糊插值推理已有许多研究,然而对多模糊变量的插值推理研究却较少。在多维规则条件下的插值推理研究方法中^[4~8],具有代表性的是MB插值推理方法^[5],它在一定的限制条件下能推理出正确的结论,但是也存在着推理结论凸性的问题。随后,文[4]对它做一定的改善,但仍存在着推理结论非正规和非凸性的问题。这些多变量模糊推理方法没有考虑到多变量之间的联系,而是独立地考虑每一变量,不论是在计算核心时还是计算斜率、插值基以及距离等时,它们按照算术平均值来计算。这样,计

*)国家科技部高新技术计划项目(2005EJ000017)、河北省科技研究与发展计划(02547015D)、河北省普通高等学校博士科研资助基金,2002(B2002118)。刘文远 博士,教授,研究方向为软计算、数据库、电子商务;张庆大 硕士研究生,研究方向为软计算、模糊推理;王宝文 副教授,硕士生导师,研究方向为软计算、模糊推理;石岩 教授,研究方向为软计算、模糊神经网络、决策与控制。

算就没有考虑到每一变量对结论的影响程度是不一样的,因而得到的结果误差是比较大的。

多变量稀疏模糊推理通常使用的规则是:

Rule 1: if x is A_{11} and A_{12} and...and A_{1n} then Y is B_1

.....

Rule m : if x is A_{m1} and A_{m2} and...and A_{mn} then Y is B_m

其中, A_j 是论域 X 上的模糊集, B_i 是论域 Y 上的模糊集, $i=1 \dots m, j=1 \dots n$ 。

然而,在现实世界中,很多问题都是受多变量影响的。在多维模糊推理中,多维变量之间不是相互独立的,它们之间有一定的联系,它们对推理结果的影响是不同的,有的变量影响大,有的变量影响较小。因此,我们需要得到变量之间的影响权值,这样更能表达它们之间的关系,使推理结果更加准确地符合实际。

规则增加权值之后,多变量稀疏模糊推理使用的规则变为:

Rule 1: if x is $A_{11}(w_{11})$ and $A_{12}(w_{12})$ and...and $A_{1n}(w_{1n})$ then Y is $B_1(CF_1)$

.....

Rule m : if x is $A_{m1}(w_{m1})$ and $A_{m2}(w_{m2})$ and...and $A_{mn}(w_{mn})$ then Y is $B_m(CF_m)$

其中, w_{ij} 表示变量的权值, CF_i 表示规则的置信度。

规则中多变量的权值如果靠人为来添加,那么工作量太大,而且不切实际。针对此问题,本文提出通过模糊神经网络进行训练,来得到模糊变量的权值;随后提出一种基于模糊神经网络加权的多变量模糊推理方法。

2 模糊 RBF 神经网络的结构和算法

在这一部分,简要介绍模糊 RBF 模糊神经网络,本文使用模糊 RBF 模糊神经网络来提取多变量的权值。

2.1 模糊 RBF 神经网络的结构

具有 n 个输入的 RBF 网络实现从 R^n 到 R 的映射 $f_r = \lambda_0 + \sum_{i=1}^m \lambda_i \theta(\|x - c_i\|)$, 其中泛函 $\theta(x)$ 具有径向对称形式,常用的函数有几种,本文采用高斯函数 $\theta(x) = \exp[-\frac{(x-c)^2}{\delta^2}]$ 。

图 2 为模糊神经网络的结构,它共有五层。

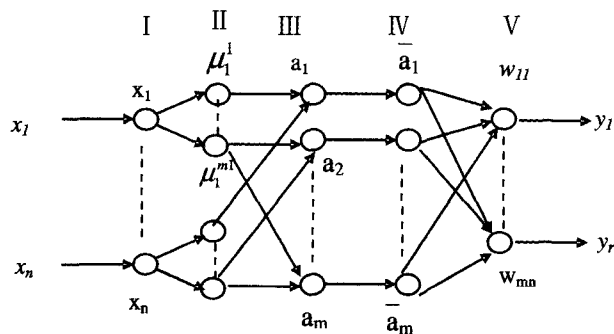


图 2 模糊 RBF 神经网络的结构

第 I 层为输入层。该层的各个结点直接参与输入向量的各分量 x_i 连接,它起着将输入值 $x = [x_1, x_2 \dots x_n]^T$ 传送到下一层的作用。该层的结点数 $N_1 = n$ 。

第 II 层的每个结点代表一个语言变量值。它的作用是计算各输入分量属于各语言变量值模糊集合的隶属函数 $\mu_i^j, \mu_i^k = \mu_{A_i^j}(x_i)$, 其中 $i=1, 2, \dots, n, j=1, 2, \dots, m_i$ 。 n 是输入向量

的维数, m_i 是 x_i 的模糊分割数。本文采用了 RBF 的高斯函数作为隶属度函数,则

$$\mu_i^j = e^{-\frac{(x_i - c_{ij})^2}{\delta_{ij}^2}} \quad (1)$$

其中, c_{ij} 和 δ_{ij}^2 分别表示隶属函数的中心和宽度。该层的结点总数 $N_2 = \sum_{i=1}^n m_i$ 。

第 III 层的每个结点代表一条模糊规则,它的作用是用来匹配模糊规则的前件,计算出规则的使用度。即

$$\alpha_j = \min\{\mu_i^1, \mu_i^2, \dots, \mu_i^m\} \quad (2)$$

该层的结点数 $N_3 = m$ 。

第 IV 层的结点数与第三层的结点数相同, $N_4 = N_3 = m$, 它所实现的是归一化计算,即

$$\bar{\alpha}_j = \frac{\alpha_j}{\sum_{i=1}^m \alpha_i}, j=1, 2, \dots, m \quad (3)$$

第 V 层是输出层,它所实现的是清晰化计算,即

$$y_i = \sum_{j=1}^r w_{ij} \bar{\alpha}_j, j=1, 2, \dots, r \quad (4)$$

2.2 模糊神经网络学习算法

假设各输入分量的模糊分割是预先确定的,那么需要学习的参数主要是最后的连接权 w_{ij} ($j=1, 2, \dots, m; i=1, 2, \dots, r$), 以及第二层的隶书函数的中心 c_{ij} 和宽度 δ_{ij} 。对于径向基函数的参数中心 c_{ij} 和宽度 δ_{ij} 采用“k-均值聚类算法”确定,这是一种在模式识别中广泛使用且性能良好的无监督学习算法,具体过程如下。

第一步:

- 1) 聚类中心 c_j 初始化 $j=1, 2, \dots, N$;
- 2) 将所有的模式按最近聚类中心分组;
- 3) 重新计算聚类中心。 $c_j = 2/M_j \sum_{x_i \in \theta_j} x_i$, M_j 表示 θ_j 中

元素数目;

- 4) 重复步骤 2), 3) 直到聚类中心 c_j 不再变化。

第二步: 聚类完成后,利用聚类中心和训练模式之间平均距离来表示 δ_j^2 :

$$\delta_j^2 = \frac{1}{M_j} \sum_{x_i \in \theta_j} (x_i - c_j)^T (x_i - c_j) \quad (5)$$

第三步: 参数 δ_j^2 确定后,对权值进行确定。设 y_{di} 和 y_i 分别表示期望输出和实际输出, $\beta > 0$ 为学习率,则网络权值 w_{ij} 按下式改变:

$$w_{ij}(k+1) = w_{ij}(k) + \beta(y_{di} - y_i) \bar{\alpha}_j x_i \quad (6)$$

这就是我们要得到的多维变量的权值。然后对它进行处理,使它保证 $\sum_{i=1}^n w_i = 1$ 。

如果我们把多变量输入到上述的模糊 RBF 神经网络中,就可以得到各个变量的权值。

3 一种基于神经网络加权的多维模糊推理方法

假定有稀疏规则:

Rule 1: if x is A_{11} and A_{12} and...and A_{1n} then Y is B_1

Rule 2: if x is A_{21} and A_{22} and...and A_{2n} then Y is B_2

观察值: A_1^* and A_2^* and...and A_n^*

结论: $B^* ???$

令 $A_1 = [A_{11}, A_{12}, \dots, A_{1n}]$, $A_2 = [A_{21}, A_{22}, \dots, A_{2n}]$, $A^* = [A_1^*, A_2^*, \dots, A_n^*]$, 其中, A_{11}, A_{21}, B_1, B_2 分别是论域 X 和 Y 中的模糊集, $i=1, \dots, n$, 因为是稀疏规则库,那么有 $A_1 < A^* < A_2$ 。

本文使用梯形隶属函数来表示模糊集,如图 3 所示。

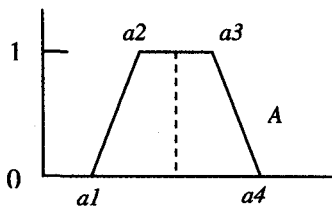


图 3 梯形隶属函数

梯形隶属函数的核值不止一个, a_2 是最小的核值, a_3 是最大的核值, 我们取这些核值的中间值称为核心。模糊集 A 的核心: $core(A) = (a_2 + a_3) / 2$

因为规则的前件有 n 个模糊变量, 所以在线性插值推理时, 需要计算 n 个插值基数, 即得到一插值基数序列 $[\lambda_1, \lambda_2, \dots, \lambda_n]$ 。令插值基

$$\lambda_i = \frac{d(core(A_{2i}), core(A^*))}{d(core(A_{2i}), core(A_{1i}))}, i=1, 2, \dots, n \quad (7)$$

把给定的相邻的规则输入到上一节提到的模糊 RBF 神经网络中, 得到规则 1 和规则 2 中多变量的权值 w_{1i} 和 w_{2i} , $i = 1 \dots n$ 。根据线性插值以及插值基数序列可以确定观察值 $A^* = [A_1^*, A_2^*, \dots, A_n^*]$ 中各变量的权值, 即

$$w_i = \lambda_i \times w_{2i} + (1 - \lambda_i) \times w_{1i} \quad (8)$$

下面分两步来求得结论 B^* :

1) 构建一条新的多变量模糊规则。

根据线性插值推理方法构造一条新的模糊规则 $A' \rightarrow B'$, 令 $A' = [A'_1, A'_2, \dots, A'_n]$ 。

首先构造 A' :

$$A'_i = \lambda_i A_{1i} + (1 - \lambda_i) A_{2i} \quad (9)$$

把 $\lambda_i = \frac{d(core(A_{2i}), core(A^*))}{d(core(A_{2i}), core(A_{1i}))}$ 带入到(9)式中, 会得到 $core$

$(A'_i) = core(A_i^*)$, $i=1, \dots, n$ 。也就是说, 得到的新规则的前件模糊集的核心与观察值的核心是相同的。

其次构造 B' 。构造 B' 需要知道规则中后件的插值基数 λ , 根据前件的插值基数序列以及前件 n 维模糊变量对后件的影响因子(即权值 w_i), 可以得到后件的插值基数为

$$\lambda = \sum_{i=1}^n w_i \lambda_i \quad (10)$$

有了插值基数 λ , 根据线性插值的原理, B' 为:

$$B' = B_1 \times \lambda + B_2 \times (1 - \lambda) \quad (11)$$

$$core(B') = core(B^*) \quad (12)$$

这样就构造了一条新的模糊规则, 原来的模糊规则就变为

Rule : if x is A'_1 and A'_2 and \dots and A'_n then Y is B'

观察值: A_1^* and A_2^* and \dots and A_n^*

结论: B^* ?

构建的新规则的前件 A' 与 A^* 是近似的, 利用 A' 与 A^* 的相似性来构造 B^* 与 B' 之间的相似性, 就可以根据相似性来得到结论 B^* 。

2) 根据相似性确定结论模糊集 B^* 。

为了表达模糊集之间几何形状的相似性, 我们定义一简单的模糊集之间相似性, 如图 4 所示。

我们把模糊集的隶属函数分为了三段, A 与 B 之间的相似性 $S_{(A,B)}$ 就分为左相似性、中间相似性和右相似性, 即

$$S(A, B) = \begin{cases} S_{(A,B)}^L = \frac{a_2 - a_1}{b_2 - b_1} \\ S_{(A,B)}^M = \frac{a_3 - a_2}{b_3 - b_2} \\ S_{(A,B)}^R = \frac{a_4 - a_3}{b_4 - b_3} \end{cases} \quad (13)$$

有了相似性就可以描述模糊集之间的关系。

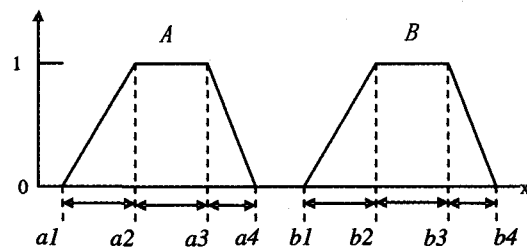


图 4 模糊集之间的相似性

A' 和与观察值 A^* 之间的相似性, 因为包含有 n 维变量, 所以得到一相似性序列 $[S_1, S_2, \dots, S_n]$, 即

$$S_i = \begin{cases} S_{(A', A^*)}^L = \frac{ai_2' - ai_1'}{ai_2^* - ai_1^*} \\ S_{(A', A^*)}^M = \frac{ai_3' - ai_2'}{ai_3^* - ai_2^*} \\ S_{(A', A^*)}^R = \frac{ai_4' - ai_3'}{ai_4^* - ai_3^*} \end{cases} \quad (14)$$

然后, 根据 A' 和 A^* 的相似程度以及多变量对结论 B^* 的影响权值 w_i , 可以计算出模糊集 B^* 与 B' 之间的相似性 S :

$$S = \begin{cases} S^L = \sum_{i=1}^n w_i S_i^L \\ S^M = \sum_{i=1}^n w_i S_i^M, i=1, \dots, n \\ S^R = \sum_{i=1}^n w_i S_i^R \end{cases} \quad (15)$$

根据式(14)、(15)可以得到模糊集 B^* 隶属函数的三段宽度:

$$\begin{cases} b_2^* - b_1^* = S^L (b_2' - b_1') \\ b_3^* - b_2^* = S^M (b_2' - b_1') \\ b_4^* - b_3^* = S^R (b_2' - b_1') \end{cases} \quad (16)$$

又因为 $core(B') = core(B^*)$, 所以

$$(a_2^* + a_3^*) / 2 = (a_2' + a_3') / 2 \quad (17)$$

根据式(16)、(17)我们就可以解得未知数 $b_1^*, b_2^*, b_3^*, b_4^*$ 的值, 这样我们就求得结论 $B^* = (b_1^*, b_2^*, b_3^*, b_4^*)$ 。

4 实例分析

在这一部分举例验证本文提出的方法的可行性, 并且与 MB 方法进行比较。

“小费问题”是国外模糊系统的教材中一个非常经典的例子, “小费问题”系统的输入为服务质量和食物质量——service and foods, 输出为小费——tip。

数据已被模糊化, 形成下面的规则:

已知规则: A_{11} and $A_{12} \rightarrow B_1$, A_{21} and $A_{22} \rightarrow B_2$

观察值: A_1^* and A_2^*

结论: $B^* ???$

$A_{11} = (1, 4, 5, 6)$, $A_{12} = (1, 5, 6, 8)$, $B_1 = (8, 9, 11, 12)$

$A_{21} = (17, 21, 22, 25)$, $A_{22} = (19, 22, 23, 28)$, $B_2 = (29, 30, 31, 32)$ 。

$A_1^* = (10, 10.5, 11, 11.5), A_2^* = (15.5, 16, 16.5, 17)$ 。

$B^* = (b_1, b_2, b_3, b_4)$

把输入变量输入到模糊 RBF 神经网络中得到规则 1 和规则 2 中变量的权值 $w_{11} = 0.3, w_{12} = 0.7; w_{21} = 0.4, w_{22} = 0.6$; 然后通过插值计算可以得到观察值中变量的权值 $w_1 = 0.34, w_2 = 0.66$ 。

从图 5 中看出, MB 方法得到的结论是非凸和非正规图形; 如图 6, 本文方法计算出的结果是 $B^* = (20.45, 20.60, 21.29, 1.52)$, 与 MB 方法相比能够保证模糊集的凸性和正规性, 推理效果更好。

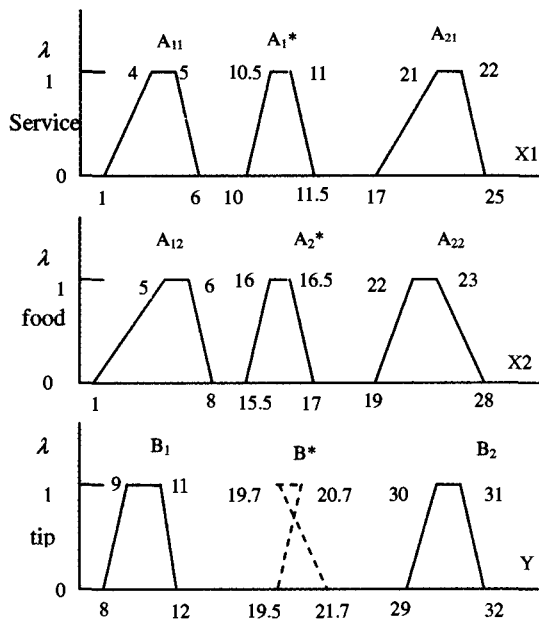


图 5 MB 方法推理结果

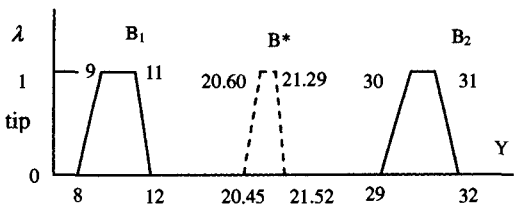


图 6 本文方法推理结果

原来的规则不变, 再给出另一观察值;

观察值: A_1^* and A_2^*

$A_1^* = (14.5, 15, 15.5, 16), A_2^* = (11, 11.5, 12, 12.5)$

结论: $B^* = (b_1, b_2, b_3, b_4) ???$

通过模糊神 RBF 神经网络计算得到变量的权值 $w_1 = 0.36, w_2 = 0.64$ 。

本文方法计算出的结果是 $B^* = (19.35, 19.50, 20.19, 20.42)$ 。

如图 7 所示, 我们看出规则没有发生变化, 观察值变了, 但是 MB 方法计算出的结果与图 5 中的结果相同, 这是因为它在计算时没有考虑到变量之间的权值, 而是按照算术平均值来计算的。然而, 如图 8, 本文由于使用模糊神经网络得到变量的权值, 所以得出的结果有所改变, 而且保证了模糊集的正规性和凸性。

结论 在多维稀疏模糊推理系统中, 许多多维模糊推理方法没有考虑到多变量之间的联系, 而是独立地考虑每一变

量, 但是多维变量之间不是相互独立的, 他们之间有一定的联系, 他们对推理结果影响是不同的。因此, 我们需要得到多变量对结论的影响权值, 本文通过模糊 RBF 神经网络进行训练, 来得到多维模糊变量的权值, 由此本文提出一种基于模糊神经网络加权的多变量模糊推理方法。首先建立一条新的模糊推理规则, 再根据模糊集之间的近似性求得结论, 该方法简单适用, 推理效果更好。

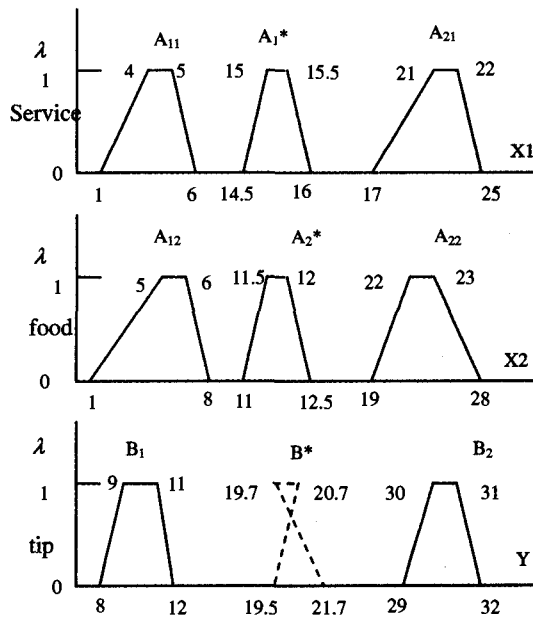


图 7 MB 方法推理结果

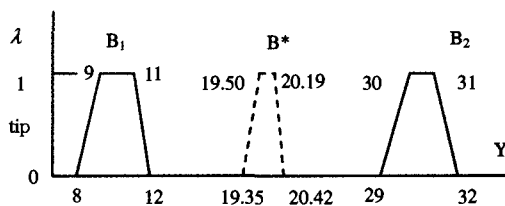


图 8 本文方法推理结果

参考文献

- 1 Kóczy L T, Hirota K. Interpolative reasoning with insufficient evidence in sparse fuzzy rules bases Inform. Sci, 1993, 71:169~201
- 2 Koczy L T, Hirota K. Approximate reasoning by linear rule interpolation and general approximation [J]. International Journal of Approximate Reasoning, 1993, 9(3):197~225
- 3 王天江, 卢正鼎, 李凡. 基于几何相似的模糊插值推理. 计算机科学, 2004, 31(9):169~171
- 4 王天江, 卢正鼎. 多变量规则的线性插值推理方法. 计算机科学, 2004, 31(6):173~176
- 5 Marsala C, Bernadette bouchon-meunier. interpolative reasoning with multi-variable rules [C]. In: Proceeding of Joint 9th IFSA World Congress and 20th NAFIPS International Conference, 2001, 7:2476~2482
- 6 Huang Z. A New Fuzzy Interpolative Reasoning Method Based on Center of Gravity. In: The IEEE International Conference on Fuzzy Systems, 2003. 25~30
- 7 Baranyi P. A generalized concept for fuzzy rule interpolation. Fuzzy systems, IEEE Transactions 2004, 12(6): 820~837
- 8 Wong K W, Tikk D, Gedeon T, et al. Fuzzy Rule Interpolation for Multidimensional Input Spaces with Applications: A Case Study. IEEE Transactions on Fuzzy Systems, 2005, 13(6): 809~819
- 9 李洪兴. 模糊控制的插值机理. 中国科学(E辑), 1998, 28(3): 259~267
- 10 Zadeh L A. Fuzzy Sets and Their Applications to Cognitive and Decision Processes[M]. In: Zadeh L A, et al. New York Academic Press, 1975. 1~39
- 11 Mizumoto M, Zimmermann H-J. Comparison of fuzzy reasoning methods[J]. Fuzzy Sets and Systems, 1982, 8(3): 253~283